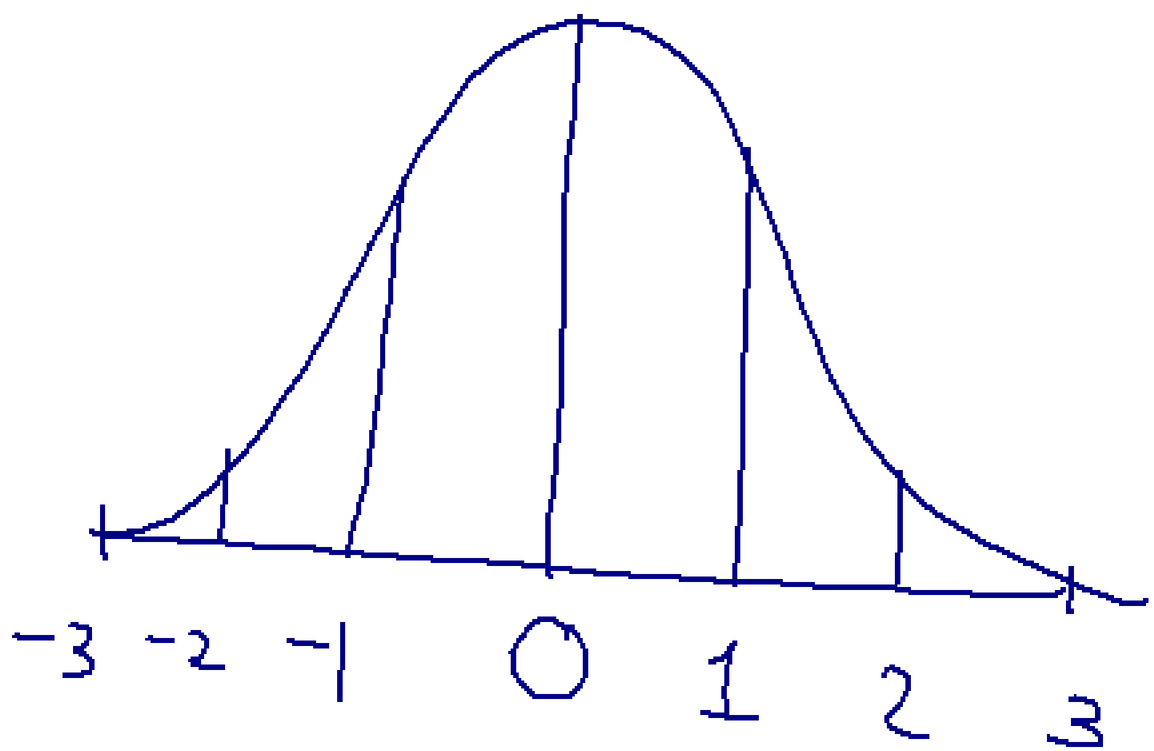
REVIEW:

Testing Proportions

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

z-distrib.



Testing Means (averages):

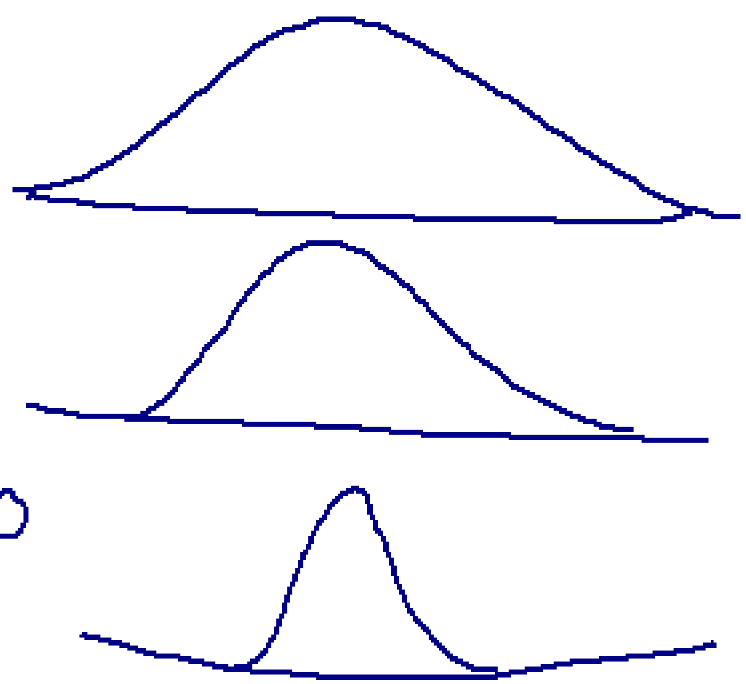
t-distrib.

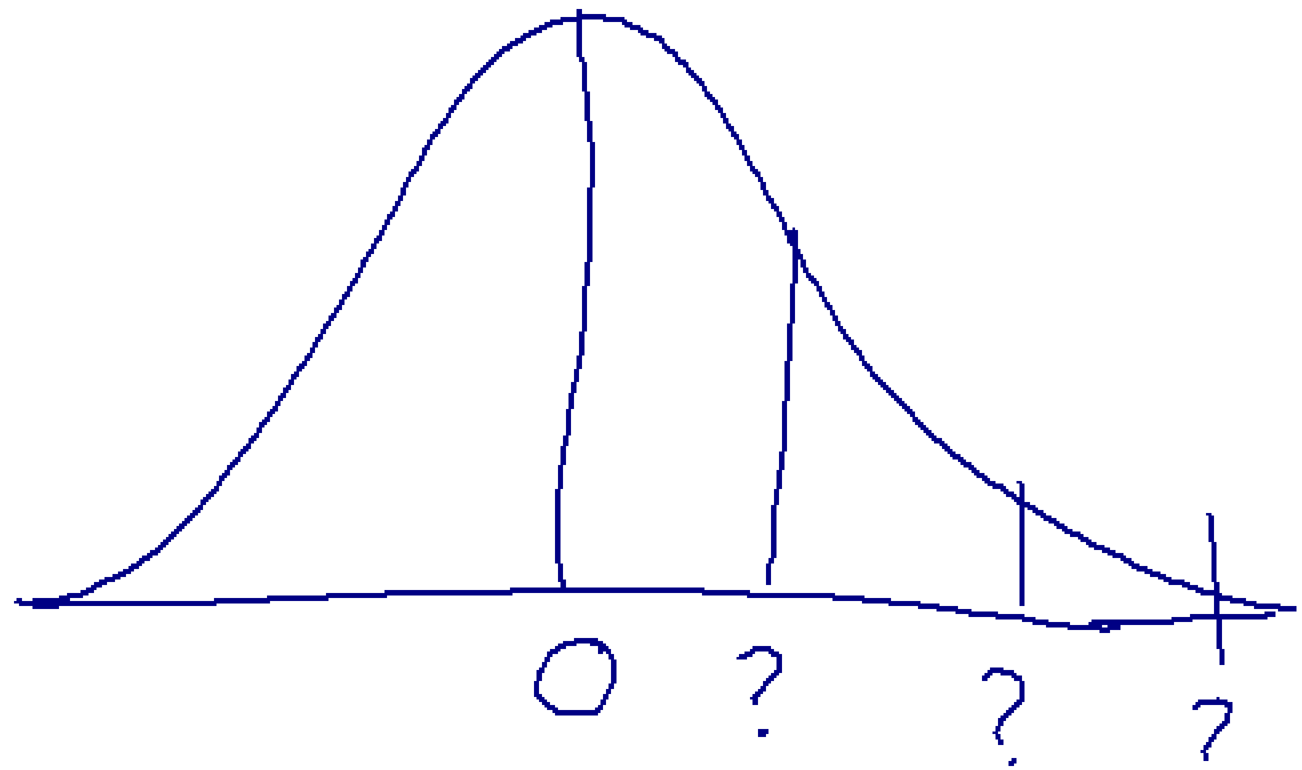
$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

n=10

n=100

n=1000





* depends on n

NEW: testing full distributions

comparing full chart to expected/claim

Ex: dice:

	1	2	3	4	5	6
Exp	10	10	10	10	10	10
Obs.	12	14	8	7	10	8

Chi-Square Goodness of Fit Test

* Testing whether... a full distribution of #'s fits an expected distrib.

HYPOTHESES: $\mu =$ $p =$

Ho: the observed sample distribution of context fits the expected distrib.

Ha: the observed sample distrib. of doesn't fit the expected distrib.

TEST STATISTIC:

Symbol: χ^2 Called: Chi-Square

Formula:

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

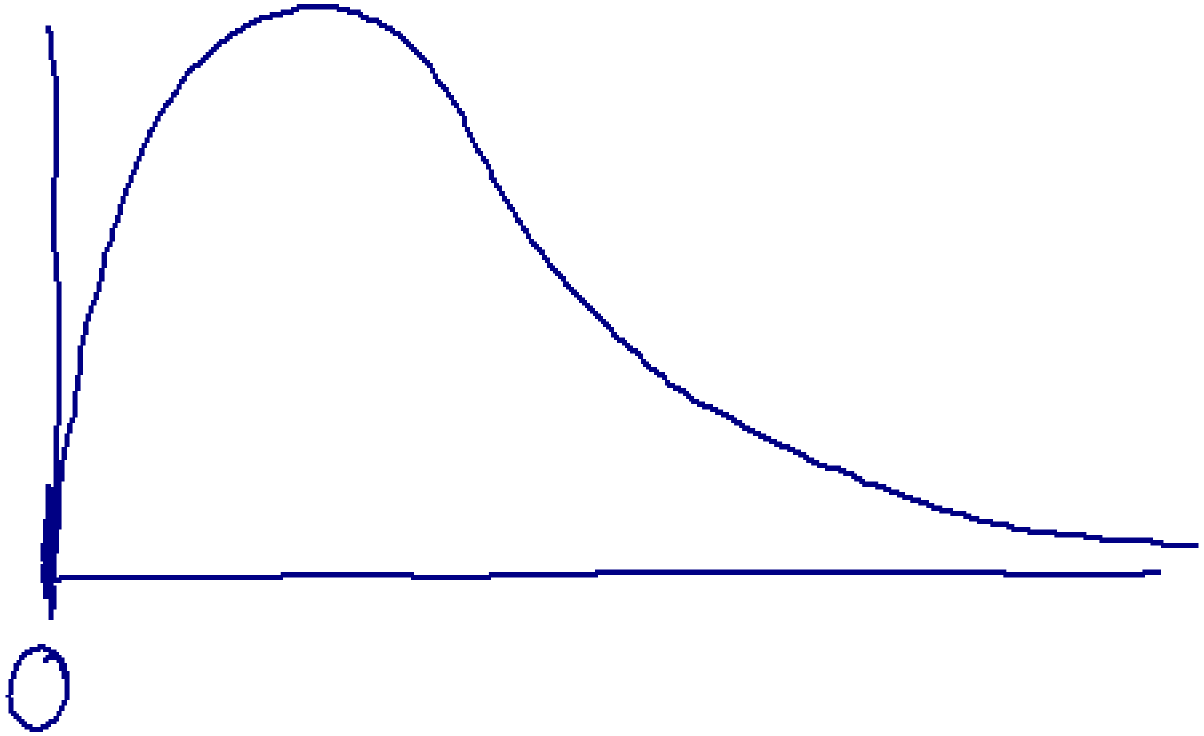
Sigma = sum

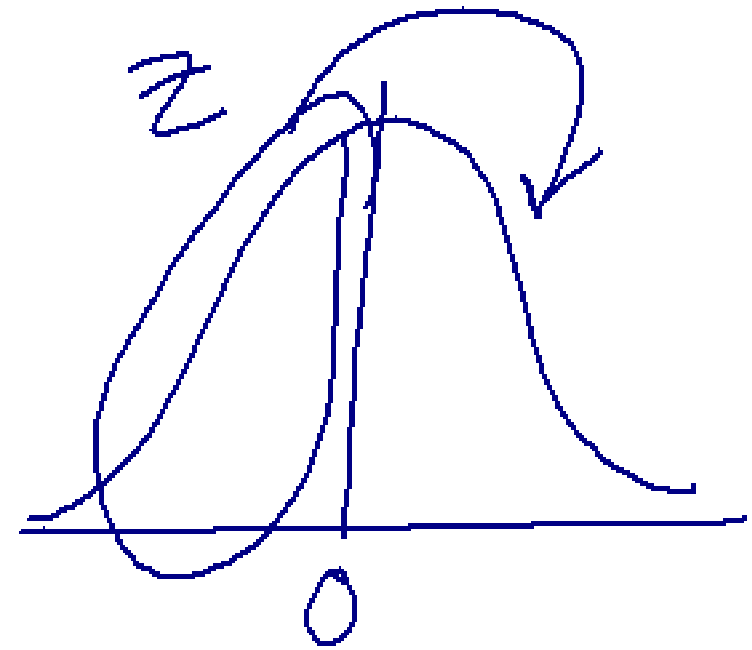
*do for each outcome

* always positive #

* big #'s

Distrib:





df = degrees of freedom

= # of outcomes − 1

Ex: dice 6 − 1 = 5

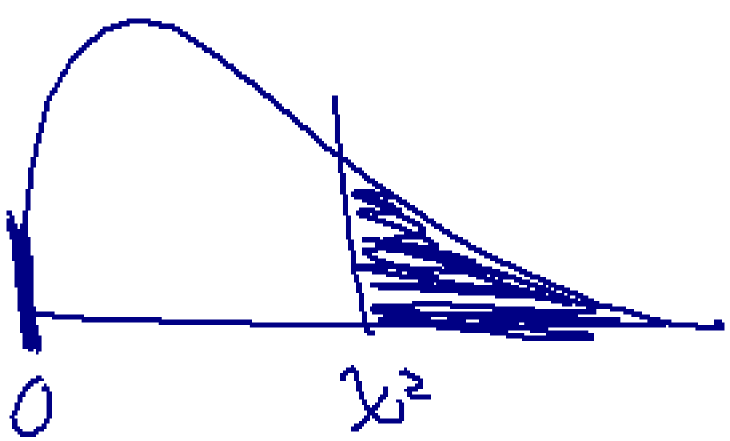
P-value

work: $P(\chi^2 > \underline{\text{test statistic}})$

E99

calculator: $\chi^2\text{cdf}(\text{lower}, \text{upper}, df)$

degrees of freedom



Always looking positive direction $\rightarrow$

Concl.

- We reject/fail to reject .....
- We have suff. evid. that (re-copy Ho or Ha)

Conditions

1. SRS
2. all expected #'s ≥ 5

#'s

$n = 300$	
1	15
2	45
3	30
4	45
5	15
6	150

H_0: the observed sample distr. of dice rolls fits the exp. distr.

Ha: the obs. sample distr. of dice rolls doesn't fit the exp. distr.

$L_3 = (L_1 - L_2)^2 / L_2$

L_1 obs	L_2 exp	Obs - Exp	$(\text{Obs} - \text{Exp})^2$	$\frac{(O-E)^2}{E}$
23	15	8	64	4.2667
50	45	5	25	0.5556
42	30	12	144	4.8
65	45	20	400	8.8889
20	15	5	25	1.667
100	150	-50	2500	16.667

$$\chi^2 = \sum \frac{(\text{obs} - \text{exp})^2}{\text{exp}}$$

$$\chi^2 = 36.844$$

$\text{sum}(L_3)$

$\chi^2 = 36.844$

$df = 6 - 1 = 5$

P-value: $P(\chi^2 > 36.844) = 6.435 \times 10^{-7}$

- reject H_0 .....

- Suff. evid. that (recopy H_a).

Ex2

$n = 124$

	Exp %	Exp # (L_2)	Obs # (L_1)
PE	25%	31	40
ME	15%	18.6	20
SE	15%	18.6	20
AE	5%	6.2	10
SpEd	20%	24.8	30
HE	10%	12.4	15
FLE	5%	6.2	10
Other	5%	6.2	9

Ho: the observed sample distrib. of education majors fits the expected distr.

Ha: the observed sample distrib. of education majors doesn't fit the exp. distr.

$$\chi^2 = \sum \frac{(obs - exp)^2}{exp} = 10.382$$

$df = 8 - 1 = 7$

$P(\chi^2 > 10.382) = 0.168$

$\alpha = 0.05$

- We fail to reject Ho b/c pvalue > $\alpha = 0.05$

- We have sufficient evidence that the observed sample distribution of education majors fits the expected distrib.

H_0: the obs samp. distr. of computer use fits the exp. distr.

H_a: the obs. samp. distr. of computer use doesn't fit the exp. distr.

$$\chi^2 = \sum \frac{(obs - exp)^2}{exp} = 0.2995$$

$df = 4 - 1 = 3$

$P(\chi^2 > 0.2995) = 0.9601$ $\quad \alpha = 0.05$

- We fail to reject H_0 b/c pvalue $> \alpha = 0.05$
- We have suff. evid that (recopy H_0)

1. Ho: the obs. samp. distr. of product displays fits the exp distr.

Ha: " " " " " " " doesn't fit " " "

$$\chi^2 = \sum \frac{(obs - exp)^2}{exp} = 8.903$$

$df = 5 - 1 = 4$

$P(\chi^2 > 8.903) = 0.0636$

- fail to reject
- re-copy Ho

② Ho: the obs. samp. distr. of coffee locations fits the exp. distr.

Ha: " " " " " " " doesn't fit " " "

$$\chi^2 = \sum \frac{(obs - exp)^2}{exp} = 3.75$$

$$df = 4 - 1 = 3$$

$$P(\chi^2 > 3.75) = 0.2893 \qquad \alpha = 0.05$$

- fail to reject Ho
- (recopy Ho)